Elo-based Similar-Strength Opponent Sampling for Multiobjective Competitive Coevolution

Sean N. Harris
BONSAI Lab
Auburn University
Auburn, Alabama, U.S.A.
snh0037@auburn.edu

Daniel R. Tauritz
BONSAI Lab
Auburn University
Auburn, Alabama, U.S.A.
dtauritz@acm.org

Abstract

Multiobjective evolution and competitive coevolution are subfields of evolutionary computation which each significantly complicate the concept of fitness. Together, these make multiobjective competitive coevolution very difficult to work with by combining multiple objective values with non-absolute fitness, preventing the use of many established techniques for improving performance in multiobjective or competitive coevolutionary algorithms. Nonetheless, multiobjective scenarios arise frequently in competitive coevolution, such as whenever coevolving agents must consider costs for their actions. This paper proposes a new evaluation method of pairing opponents with similar skill levels in each objective, so that evaluations more efficiently distinguish the performance of similar individuals. This is enabled through the use of per-objective Elo ratings as a surrogate fitness function that prevents bias against individuals assigned stronger opponents. Ratings can further be assigned for asymmetric, non-zero-sum objectives such as cost, allowing individuals to be paired with opponents that incidentally challenge those asymmetric objectives. Mixed results are presented, showing significant benefits from pairing similar opponents, but finding that the use of Elo rating instead of raw fitness harms evolution. A novel statistical test for comparing multiobjective coevolutionary algorithms is also introduced.

CCS Concepts

• **Computing methodologies** → **Genetic algorithms**; **Adversarial learning**; • **Theory of computation** → **Evolutionary algorithms**; • **Applied computing** → **Multi-criterion optimization and decision-making**.

Keywords

Evolutionary Algorithms, Competitive Coevolution, Competitive Co-evolution, Multiobjective Evolutionary Algorithms, Elo Rating System

ACM Reference Format:
Sean N. Harris and Daniel R. Tauritz. 2021. Elo-based Similar-Strength Opponent Sampling for Multiobjective Competitive Coevolution. In *2021 Genetic and Evolutionary Computation Conference Companion (GECCO '21 Companion), July 10–14, 2021, Lille,France.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3449726.3459559

GECCO '21 Companion, July 10–14, 2021, Lille, France

ACM ISBN 978-1-4503-8351-6/21/07.
https://doi.org/10.1145/3449726.3459559

1 Introduction

This paper introduces a novel method of sampling opponents in competitive coevolutionary algorithms by pairing solutions with opponents found to have similar skill levels in past evaluations (Similar-Strength Opponent Sampling, SSOS). This is intended to reduce the number of needed evaluations per generation by more efficiently distinguishing the skill levels of similar opponents. Biasing opponent selection biases fitness values as well, so we apply the Elo rating system [3] as a surrogate to get an absolute fitness measure. For each unique objective, each individual is assigned a separate Elo rating. Even in asymmetric games, individuals still receive ratings for opponents' objectives, representing their ability to challenge that objective despite lacking selective pressure to do so. As a result, Elo-Based Similar-Strength Opponent Sampling (EBSSOS) can be applied to games that other methods are inapplicable to due to their focus on single-objective, zero-sum games.

2 Elo Rating System

The Elo rating system was originally developed by Arpad Elo [3] for rating the skill level of chess players. We employ a method provided by Elo to calculate ratings for a group of unrated players, referred to as "the method of successive approximations". Given a previous rating estimate for the population, each individual's rating estimate can be updated by the following:

$$R_p = \gamma \cdot R_c + 200 \cdot \log_{\sqrt{10}}(\frac{P}{1-P}) \tag{1}$$

where R_p is the new rating estimate, R_c is the average of the rating estimates of all that individual's opponents, P is the average of the scores that individual has achieved (normalized to $[0,1]$), and γ is a decay value added to prevent divergence for low numbers of games, set to 0.9. We repeat these updates until the average change in ratings between iterations is below a threshold of 0.1. These parameters were found to produce stable ratings quickly.

3 Problem Domain

We evaluate our algorithm in a two-agent predator-prey environment. Predator-Prey games are common for testing competitive coevolutionary algorithms, allowing complex behaviors despite their simplicity [1, 5]. Our environment consists of agents which move at fixed speeds, and are controlled by selecting an angle to move for each timestep. Agents are circular with a radius of 0.1, where 1.0 is the radius of the circular world. If the predator overlaps the prey, or the prey survives for 200 time steps, the game ends. Moves directed outside the outer edge are projected towards the center of the world to the nearest valid location. Agents start at opposite sides, half-way to the edge. We use a predator speed of 0.06, and a prey speed of 0.10, selected through tuning for fairness.

Predator-Prey games typically have a single zero-sum objective: the duration that the prey survives. In order to test multiobjective evolution, the predator and prey are each given a region of the map that fulfills a non-zero-sum "comfort" objective: the prey must minimize its average distance from the center of the world, and the predator must minimize its average distance to the edges. These objectives are chosen to conflict with the dominant strategies for the pursuit-evasion objective. Agents have no incentive to degrade their opponents' comfort objective. Despite this, SSOS assigns predators and prey opponents which incidentally challenge these objectives.

4 Methodology

Independent Elo ratings are maintained for each objective, per individual, including opponents' objectives. Evaluations are performed in several rounds, where every individual is assigned one opponent per round. Opponents are chosen to be close in Elo rating, maximizing the information about skill differences gained from each evaluation, and providing opponents that incidentally challenge individuals' asymmetric objectives. We aim to minimize the sum of squared distances in per-objective Elo ratings between paired individuals. This is accomplished by a hill-climber that randomly swaps pairings and reverses those that increase cost, terminating after 200 iterations without improvement. After a round of evaluations has been completed, the recorded objectives are normalized to $[0, 1]$ on the range of objective values yet observed. From these values, their Elo ratings are recalculated through the method of successive approximations, with an initial estimate of 0. These new ratings include transitive relationships encoded in their opponents' Elo rating from past evaluations. After all rounds are completed, the final Elo ratings are returned as the individuals' fitness. We hypothesize that this will decrease the amount of evaluations per generation necessary to achieve healthy coevolution.

4.1 Evolutionary Design

Our predator and prey agents are strongly-typed genetic programming trees [4] which return an angle to move at for each timestep, using angle and distance types. As sensors, agents have the distance and angle to the center of the world, to the edge, and to their opponent, as well as the past move angles made by them and their opponent, enabling predictive behavior. These sensors are combined through basic arithmetic functions, scalar multiplication of angles with distances, and a function which returns one of two angles based on which of two distances is larger. Agents are evolved through NSGA-II [2] for 50 generations, using a (50+150) population model, a mutation rate of 50%, and trees initialized with ramped half-and-half to a height between 3 and 7.

4.2 Comparison Methodology

Given two runs of different multiobjective coevolutionary algorithms, we sample the k top individuals from each run's population(s), sorting by NSGA-II-styled crowded comparison. These individuals are entered into a round-robin tournament. The two runs are each scored by the fraction of their own individuals in the newly-calculated Pareto fronts for each population. A run is considered superior if it makes up the majority of the new Pareto fronts, indicating that its best solutions mostly dominate its opponent's. Statistical analysis is performed over 100 runs, where each run from the first configuration is compared against each run in the second configuration, and their resulting scores summed per-run. Two-tailed F-tests and t-tests are then be used to measure whether one configuration produces a significantly higher average share of non-dominated solutions against runs from its opponent.

Paired Elo vs. Unpaired Elo 5 Opponents per Individual	Means	71.6% vs. 28.4%
	Best	**Paired Elo**
Paired Elo vs. No Elo 5 Opponents per Individual	Means	30.9% vs. 69.1%
	Best	**No Elo**
Unpaired Elo vs. No Elo 5 Opponents per Individual	Means	15.2% vs. 84.8%
	Best	**No Elo**
Paired Elo vs. Unpaired Elo 10 Opponents per Individual	Means	80.7% vs. 19.3%
	Best	**Paired Elo**
Paired Elo vs. No Elo 10 Opponents per Individual	Means	41.2% vs. 58.8%
	Best	**No Elo**
Unpaired Elo vs. No Elo 10 Opponents per Individual	Means	14.0% vs. 86.0%
	Best	**No Elo**
Paired Elo vs. Unpaired Elo 20 Opponents per Individual	Means	79.8% vs. 20.2%
	Best	**Paired Elo**
Paired Elo vs. No Elo 20 Opponents per Individual	Means	43.4% vs. 56.6%
	Best	**No Elo**
Unpaired Elo vs. No Elo 20 Opponents per Individual	Means	17.3% vs. 82.7%
	Best	**No Elo**

Table 1: Experimental results displaying the average fraction of dominant strategies in pairwise comparisons.

5 Results and Conclusions

We compare EBSSOS with K rounds of evaluations ("Paired Elo") against two other methods of sampling: K-random opponent sampling ("No Elo") and K-random sampling with Elo rating as fitness ("Unpaired Elo"), with $K = 5$, 10, and 20. "Unpaired Elo" allows the effects of SSOS to be separated from the direct effects of using Elo rating over raw fitness. Results are shown in Table 1. All results are significant ($p < 0.05$). "Paired Elo" (EBSSOS) was found to significantly underperform "No Elo". However, the remaining results indicate a complex relationship. Compared to "Unpaired Elo", "Paired Elo" produced a significant improvement, demonstrating that the SSOS component of this algorithm is itself effective. Comparing "Unpaired Elo" to "No Elo" shows that using Elo rating as a surrogate fitness function is harmful enough to cancel out the benefits of SSOS. We suggest that Elo's transitive model does not sufficiently model relative agent performances. However, SSOS isn't reliant upon Elo rating specifically, it only needs an unbiased method of scoring individuals after biased pairings. As a result, techniques besides Elo that more effectively model these relationships might provide more useful inferences of fitness.

References

[1] Dave Cliff and Geoffrey P. Miller. 1995. Co-evolution of pursuit and evasion II: Simulation Methods and results.

[2] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6, 2 (April 2002), 182–197. https://doi.org/10.1109/4235.996017

[3] Arpad E. Elo. 1978. *The rating of chessplayers, past and present.* Arco Publishing.

[4] David J. Montana. 1995. Strongly Typed Genetic Programming. *Evol. Comput.* 3, 2 (June 1995), 199–230. https://doi.org/10.1162/evco.1995.3.2.199

[5] Craig W. Reynolds. 1994. Competition, coevolution and the game of tag. In *Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems.* 59–69.